

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 1, 2001, 16:10:43 ; Search time 32.86 Seconds

(without alignments)
136.380 Million cell updates/sec

Title: US-09-331-631A-24_COPY_29_94

Perfect score: 382

Sequence: 1 HDEDDRRGSHSLQCVQRC.....EQEEDGGRGHWGEGERE 66

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 195891 seqs, 67900655 residues

Total number of hits satisfying chosen parameters: 195891

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	382	100.0	637	2 S35221	globulin Bgl prec
2	138	36.1	582	2 B53234	vicilin-like stora
3	135	35.3	407	2 T02258	globulinl - maize
4	133	34.8	540	2 S21825	vicilin-like stora
5	133	34.8	573	2 A53234	globulin-1S, GLBIS
6	129.5	33.9	122	2 C53234	globulin-10 - maize
7	127	33.2	236	2 T01662	globulin-1 - maize
8	93	24.3	509	2 S08059	alpha-globulin typ
9	92	24.1	625	2 A34615	proflaegrin - rat
10	90	23.6	588	1 FWCNMB	alpha-globulin B P
11	85	22.3	639	2 B24810	beta-conglycinin a
12	83.5	21.9	605	2 S06398	alpha-globulin typ
13	76.5	20.0	338	2 S04321	legumin B (clone p
14	76.5	20.0	566	2 T06453	probable legumin B
15	74.5	19.5	1655	2 T13998	gene mastermind pr
16	73.5	19.2	810	2 T44430	protein FV100 (imp
17	73	19.1	605	1 FWSYTA	beta-conglycinin a
18	73	19.1	605	2 S20007	beta-conglycinin a
19	73	19.1	613	2 S27770	hypothetical prote
20	72	18.8	566	2 S22477	vicilin precursor
21	72	18.8	852	2 A34373	histidine-rich cal
22	71	18.6	564	2 S37241	legumin B fava b
23	70.5	18.5	185	2 S71512	hypothetical prote
24	70	18.3	483	2 T06459	62K sucrose-bindin
25	70	18.3	559	2 S52941	period protein (cl
26	70	18.3	560	2 S52943	period protein (cl
27	70	18.3	1122	2 B26427	period clock prote
28	70	18.3	1127	2 A25018	circadian rhythm p
29	70	18.3	1130	2 T30251	repetin - mouse

30	70	18.3	1176	2 C26427	period clock prote
31	70	18.3	1218	2 A26427	period clock prote
32	70	18.3	1218	2 A26588	period clock prote
33	70	18.3	1549	1 A40691	trichohyalin - she
34	69.5	18.2	250	2 B35026	filaggrin B - mus
35	69.5	18.2	255	2 A35026	filaggrin A - mus
36	69.5	18.2	1208	2 S17286	period clock prote
37	69	18.1	558	2 S52951	period protein (cl
38	69	18.1	558	2 S52944	period protein (cl
39	68.5	17.9	654	2 S11049	potassium channel
40	68.5	17.9	654	2 E39113	potassium channel
41	67.5	17.7	254	2 A31488	filaggrin - mouse
42	67.5	17.7	313	2 A28444	filaggrin precursor
43	67.5	17.7	590	2 A40437	glutamic acid-rich
44	67	17.5	228	2 S19132	rib25 protein - ri
45	67	17.5	516	1 FWCNMB	beta-globulin B pr

ALIGNMENTS

```

RESULT 1
S35221
globulin Bgl precursor - barley
C:Species: Hordeum vulgare (barley)
C:Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 21-Jul-2000
C:Accession: S35221
R:Heck, G.R.; Chamberlain, A.K.; Ho, T.H.D.
Mol. Gen. Genet. 239, 209-218, 1993
A:Title: Barley embryo globulin 1 gene, Bgl1: characterization of cDNA, chromosome ma
A:Reference number: S35221; MUID:93287988
A:Accession: S35221
A:Molecule type: mRNA
A:Residues: 1-637 <HEC>
A:Cross-references: EMBL:M64372; NID:9167003; PIDN:AAA32936.1; PID:9167004
C:Genetics:
A:Gene: Bgl1
A:Map position: 4
C:Superfamily: glycinin
C:Keywords: glycoprotein
P:174-190/Product: globulin Bgl1 #status predicted <MA>

Query Match 100.0%; Score 382; DB 2; Length 637;
Best Local Similarity 100.0%; Pred. No. 18e-32;
Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HDEDDRRGSHSLQCVQRCQERPRYSHARCVOECRDQOQHGHRHQEEDGGRGHWG 60
      |||||
Db 29 HDEDDRRGSHSLQCVQRCQERPRYSHARCVOECRDQOQHGHRHQEEDGGRGHWG 88

QY 61 EGEREE 66
      |||||
Db 89 EGEREE 94

RESULT 2
B53234
vicilin-like storage protein Gb1-L, embryo - maize
N:Alternate names: globulin-IL
C:Species: Zea mays (maize)
C:Date: 02-May-1994 #sequence_revision 18-Nov-1994 #text_change 24-Nov-1999
C:Accession: B53234; S21824
R:Belanger, F.C.; Kriz, A.L.
Genetics 129, 863-872, 1991
A:Title: Molecular basis for allelic polymorphism of the maize globulin-1 gene.
A:Reference number: A53234; MUID:92090707
A:Accession: B53234
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-582 <BEL>
A:Cross-references: EMBL:X59083; NID:922283; PIDN:CAA41809.1; PID:922284
A:Experimental source: inbred line W64A6

```


Query Match 52: 33.9% Score 129.5: DB 2: Length 122:
 Best Local Similarity 37.9% Pred. No. 7,8e-07:
 Matches 22: Conservative 13: Mismatches 18: Indels 5: Gaps 2:

QY 2 DDEDRCRGHSIQQCVQRCRQREPRYSARQYQEDRDD----QQQGRGEQEEQGRG 55
 |||||
 Db 26 DNNNNHNGKSGQGVRC-EDRPNNQRDRCEQGRREERKREKREKREKREKREKREKREK 82
 |||||

RESULT 7
T01662 globulin-1 - maize (fragment)
C:Species: Zea mays (maize)
C:Date: 19-Feb-1999 #sequence
C:Accession: T01662 #lexn_change 11-Jan-2000
R:Hilton, H.; Gaut, B.S.
Genetics 150, 863-872, 1998
A:Title: Speciation and domestication in maize and its wild relatives: evidence from the
A:Reference number: Z14586; MUID:98425537
A:Accession: T01662
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-236 <Hill>
A:Cross-references: EMBL:AF064222; NID:q3414836; PIDN:AA031465.1; PID:93414837
A:Experimental source: subspecies parviglumis
C:Genetics:
A:Introns: 166/1; 224/3
A:Superfamily: glycinnin

Query Match	33.2%	Score 127;	DB 2;	Length 236;
Best Local Similarity	37.5%	Pred. NC. 2.6e-06;		
Matches 24;	Conservative 14;	Mismatches 20;	Indels 6;	Gaps 3;

[illegible]

RESULT 8
S08059
alpha-globulin type B precursor (random 1) - upland cotton (fragment)
N:Alternate names: seed storage protein
C:Species: Gossypium hirsutum (upland cotton)
C:Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 30-Sep-1993
C:Accession: S08059
R:Chlan, C.A.; Borroto, K.; Kamalay, J.A.; Dure III, L.
Plant Mol. Biol. 9, 533-546, 1987
A:Title: Developmental biochemistry of cottonseed embryogenesis and germination. XIX. Seed
A:Reference number: S06398
A:Accession: S08059
A:Status: not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-509 <CHL>
C:Superfamily: g1ylc1n1n

Query Match	24.3%	Score 93;	DB 2;	Length 509;
Best Local Similarity	29.4%	Pred. NO. 0.018;		
Matches 20; Conservative	16;	Mismatches 24;	Indels 8;	Gaps 3;

DQ 3 DEDRRGGHSIQCCVACRCRGRER-YSHARCVQECRDDQQH--GRHQEEEDGGRGW 58
::: | | | | : ||| | : | | |
Db 33 EQEQQQSQRFQEQCQHCHQEQRPEKKQCVRCEKREKYENPMRGEREFEAEETE- 90

QY	59	HGEGEREE	66
		: :	
Db	91	--EGEQEQ	96

RESULT 9
A34615
profilaggrin - rat (fragment)
C.Species: Rattus norvegicus (Norway rat)
C.Date: 29-Jun-1990 #sequence_revision 09-Oct-1992 #text_change 04-Feb-2000
C.Accession: A34615

A:Title: Filaggrin, an intermediate filament-associated protein: structural and functional properties
A:Reference number: A34615; MUID:90274870

A;Status: preliminary
A;Molecule type: mRNA
A;Positives: 1-635 (HAV)

A:Cross-references: GB:M21759; NID:g204143; PIDN:AAA1161.1; PID:g204144
A:Note: the authors translated the codon GAA for residue 568 as Gln
A:Superfamily: unassigned calmodulin-related proteins; calmodulin repeat homology
C:Keywords: EF hand; epidermis

Query Match	24.1%	Score	92:	DB	2:	Length	625:
Best Local Similarity	35.1%	Pred. NO.	0.028:				
Matches	26:	Conservative	7:	Mismatches	21:	Indels	20:
						Gaps	3:

```

QY      2 DDEDDRGGHSLQCCYQRCRQGEPRYSIARCVQCRDDQQ-----HGR 45
      :||: || || || :||: || || ||
Db      319 EDESDAQCGHQQQQQQR-QQCRQQRQHE--GREHEQQPESQGRQCGSGRTQNAAR 374

```

```
QY      46 HEQEEOGRGRGNH 59
      III: : IIII
Db      375 HEQDSDSTRQGRSH 388
```

RESULT 10

alpha-globulin B precursor (clone C72) - upland cotton
 N:Alternate names: seed storage protein; villin precursor
 C:Species: *Gossypium hirsutum* (upland cotton)
 C:Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 16-Jul-1999
 C:Accession: A30838; S06311
 R:Chian, C.A.; Pyle, J.B.; Legocki, A.B.; Dure III, L.
 Plant Mol. Biol. 7, 475-489, 1986
 A:Title: Developmental biochemistry of cottonseed embryogenesis and germination XVIII
 A:Reference number: A30838
 A:Accession: A30838
 A:Molecule type: mRNA
 A:Residues: 1-588 <CHL>
 A:Cross-references: GB:M6891; NID:g167374; PID:AAA35071.1; PID:g167375
 A:Experimental source: var. Coker 201
 R:Chian, C.A.; Botto, K.; Kamalay, J.A.; Dure III, L.
 Plant Mol. Biol. 9, 533-546, 1987
 A:Title: Developmental biochemistry of cottonseed embryogenesis and germination. XIX.
 A:Reference number: S06398

A:Status: not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-81 <CH2>
C:Comment: This is a seed storage protein.

C:Keywords: glycoprotein; seed; storage protein
F:1-25/Domain: signal sequence #status predicted <SIG>
F:26-588/Product: alpha-globulin storage protein #status predicted <MAT>
F:417/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match	23.6%	Score 90:	DB 1;	Length 588;
Best Local Similarity	31.2%;	Pred. No.	0.043;	
Matches 20;	Conservative 14;	Mismatches 28;	Indels 2;	Gaps 2;

QY 3 DEDDRGGSHLQQCYKRCRDERPRYSH-ARCYQECRRDQQH-GRHQEEEEGGKRGCMWG 60
:: :: | : | | : | : ||| : | : |
Dd 114 EQEQQQSQRFQEQQHCHQEQQRPEKKQCVRECRERYQENPMWREREAEAEFEETEEG 173

